

SOME PROPERTIES OF POPULATIONS WITH MULTILOCUS HOMOMORPHIC GAMETOPHYTIC INCOMPATIBILITY SYSTEMS

DEBORAH CHARLESWORTH

School of Biological Sciences, University of Sussex, Brighton BN1 9QG, Sussex, England

Received 13.iii.79

SUMMARY

Equations for the changes in genotype frequencies in two- and three-locus incompatibility systems are given. Assuming equal allele numbers at all loci, and equal frequencies of all the gamete types, as would be expected in the equilibrium state, the equilibrium ratios of genotypes heterozygous at different numbers of loci are obtained. Some genetic and evolutionary implications of these results are discussed.

1. INTRODUCTION

In the two-locus incompatibility system found in grasses (Lundqvist, 1954, 1961, 1962*a*, 1965; Hayman 1956; Murray 1974), the two loci (usually called S and Z) act so as to generate an incompatibility reaction only when both the alleles carried in the pollen grain are matched in the stigma which it has pollinated. The number of different pollen types is equal to the number of gamete types that can be generated by two loci, and is thus $n_1 n_2$ where n_1 and n_2 are the numbers of alleles at the S and Z loci. Each pollen type is incompatible with every genotype that can generate the gamete type corresponding to that carried in the pollen. Since double homozygotes cannot be formed, owing to the incompatibility system, the total number of genotypes is $n_1 n_2 (n_1 n_2 - 1) / 2$, treating each double heterozygote as two genotypes ("coupling" and "repulsion"). The number of singly and doubly heterozygous types is also easily calculated, but even at equilibrium with all single heterozygotes having the same frequency and all double heterozygotes equal in frequency, the relative frequencies of single versus double heterozygotes are not immediately obvious. Lundqvist (1962*b*) gives formulae for the frequencies of these two classes of genotype, in terms of the numbers of alleles at the two loci, but their derivation is not given. In trying to derive these results, I have obtained rather different ones. The problem is quite complex and it seems that a general solution is possible only when $n_1 = n_2$.

2. TWO INCOMPATIBILITY LOCI

Given the frequency of each genotype in one generation, it is easy to write down the frequencies in the next generation. Assuming that the two loci are unlinked, they are:

(i) *Single heterozygotes*

$$f_{ij} = \left\{ g_i \left[\sum_a \frac{f_{ja}}{2} + \sum_b \frac{f_{jb}}{4} \right] + g_j \left[\sum_c \frac{f_{ic}}{2} + \sum_d \frac{f_{id}}{4} \right] \right\} / W$$

where the f_{ja} are the frequencies of single heterozygotes that are compatible with pollen type i and that can generate gamete type j (and f_{ic} is the converse),

the f_{jb} are the frequencies of double heterozygotes compatible with pollen type i that can generate j (and f_{id} is the converse),

g_i is the frequency of gamete type i .

and W is a normalising factor which enters in because some genotypes do not contribute to any particular genotype, due to incompatibility.

(ii) *Double heterozygotes*

$$f_{kl} = \left\{ g_k \left[\sum_e \frac{f_{le}}{2} + \sum_f \frac{f_{lf}}{4} \right] + g_l \left[\sum_g \frac{f_{kg}}{2} + \sum_h \frac{f_{kh}}{4} \right] \right\} / W$$

Now if we assume that all the gametes are present at the same frequency, and write p_1 and p_2 for the frequencies of individual single and double heterozygote genotypes respectively, these equations become

$$p_1 = (g_i + g_j) \left(p_1 \frac{a}{2} + p_2 \frac{b}{4} \right) / W \quad (1)$$

$$p_2 = (g_k + g_l) \left(p_1 \frac{e}{2} + p_2 \frac{f}{4} \right) / W$$

where a is simply the number of single heterozygotes that are compatible with pollen type i and capable of generating gamete type j , and analogous interpretations apply to b , e , and f .

We then have

$$\alpha = \frac{p_1}{p_2} = \frac{2ap_1 + bp_2}{2ep_1 + fp_2}$$

which gives a quadratic equation for α . The only problem is to determine the values of the constants a , b , e and f . These depend on the numbers of alleles. Before discussing this, it may be noted that linkage of the S and Z loci does not affect the equations written above. This is obvious for the contributions from single heterozygotes but is also true for double heterozygotes if it is assumed that all gametes are equally frequent. This is because, for each gamete type j to be generated, there is an equal number of genotypes that contain j as a non-recombinant gamete, and that generate j by recombination. If the genotype frequencies are equal, the terms b and f are thus each multiplied by $\left(\frac{1-R}{4} + \frac{R}{4} \right)$ where R is the recombination fraction, and this factor is simply $\frac{1}{4}$ as for unlinked genes.

To determine the values of a , b , e and f , assume that there are n alleles at each locus. (The reason why unequal numbers of alleles cannot be dealt with is explained in the paragraph where the determination of b is

described.) Thus there are n^2 different types of gametes, and $\frac{1}{2}n^2(n^2 - 1)$ genotypes. The number of double heterozygote types is

$$\frac{1}{2}n^2(n-1)^2,$$

and the number of single heterozygotes is

$$n^2(n-1).$$

Any particular pollen type, say i , is incompatible with any genotype that can generate gamete i . Of the $2(n-1)$ single heterozygote genotypes containing j , one will also contain i , so there are $2n-3$ single heterozygous types compatible with pollen type i and capable of generating gamete type j . Thus we have $a = 2n-3$.

We next consider the evaluation of e , the number of single heterozygotes (k/j) that are compatible with i pollen, and can generate gamete j , which differs from i at both loci so that a double heterozygote is generated. Since j differs from i at both the S and Z loci, and k carries the same allele as j at one or other locus, it is impossible for the genotype j/k to match the gamete i at both loci. Hence i pollen is compatible with all single heterozygotes that can generate a gamete j that with i will form a double heterozygote. Thus

$$e = 2(n-1).$$

Now consider the progeny from double heterozygotes. Each double heterozygote produces four gamete types, and will inhibit pollen grains carrying any of those types. The number of double heterozygotes that are capable of producing any given gamete type j is thus $2(n-1)^2$. We need to know how many of these are incompatible with some particular pollen type i . To determine b , we must assume that the progeny type i/j is a single heterozygote. We can write this in terms of the individual S and

Z loci as $\frac{S_i Z_i}{S_j Z_j}$, where either

$$S_i = S_j, Z_i \neq Z_j, \quad \text{or} \quad S_i \neq S_j, Z_i = Z_j.$$

If $(S_j Z_j)$ is a non-recombinant gamete, the genotype that produced it ($j/k = \frac{S_j Z_j}{S_k Z_k}$) will be incompatible with i pollen only if gamete type i can be produced by recombination between j and k . (The possibility $k = i$ is ruled out by the assumption that i/j is a single heterozygote.) If the locus at which i/j is heterozygous is the S locus, this implies that $Z_j = Z_i$, and

incompatibility of j/k with i pollen can result only if $S_k = S_i$. Thus $j/k \equiv \frac{S_i Z_k}{S_j Z_i}$

in this case. If the number of alleles at the Z locus is n_1 , there are $n_2 - 1$ such genotypes. If the locus at which i/j is heterozygous is the Z locus, there are similarly $n_1 - 1$ genotypes which can produce j as a non-recombinant gamete, and are incompatible with i (where n_1 is the number of alleles at the S locus). Only if we assume $n_1 = n_2$ can we obtain the number of genotypes that meet the requirements specified, as an expression that does not depend on the particular genotypes in question.* This explains

* Alternatively, if $n_1 \neq n_2$, the basic equations, e.g. (1), must be replaced by a set of three equations, which yield two simultaneous quadratic equations, which cannot be solved explicitly.

the choice of $n_1 = n_2 = n$ above, and in what follows. In that case, we have (when j is a non-recombinant gamete) the result that $n-1$ double heterozygotes are excluded from contributing to the formation of singly heterozygous i/j progeny. Similarly, if j is a recombinant gamete, the contribution from the double heterozygote i/k , where k is the gamete that with i can generate j by recombination, must be subtracted because this type is incompatible with i pollen. There are $n-1$ such genotypes. Since no other double heterozygotes that can produce j are also incompatible with i , we have

$$\begin{aligned} b &= 2(n-1)^2 - 2(n-1) \\ &= 2(n-1)(n-2) \end{aligned}$$

Finally, we require f , the contribution of double heterozygotes to the frequency of i/j , where this type is itself a double heterozygote. If j is a parental gamete produced by the genotype j/k , there are two ways in which incompatibility with i can arise. First, $k = i$ (one such genotype exists). Second, k is such that i is produced by recombination. But this is impossible because j differs from i at both the S and Z loci. Similarly, if j is a recombinant gamete, produced by a double heterozygote k/l , then the reciprocal recombinant could be i , but this is the only way in which such a genotype could be incompatible with i : the possibilities $k = i$ or $l = i$ are ruled out, because it would then be impossible for k/l to generate j by recombination, where j differs from i at both the S and Z loci. Thus, of the double heterozygotes that can generate suitable js , two are incompatible with i . We therefore have:

$$f = 2(n-1)^2 - 2$$

These values of the coefficients lead to the equation:

$$\alpha = \frac{\sqrt{(n+1)^2 - 8} - (n-3)}{4} \quad (2)$$

Table 1 shows some calculated frequencies using equations (1). These have also been checked by direct computer modelling of the detailed work-

TABLE I
Frequencies of single and double heterozygotes in the two-locus system, from equations (1)

Number of alleles at each locus (n)	Single heterozygotes		Double heterozygotes	
	Individual genotype frequencies (p_1)	Aggregate frequency	Individual genotype frequencies (p_2)	Aggregate frequency
2	0.125	0.5	0.25	0.5
3	0.02301	0.414	0.03254	0.586
4	0.00713	0.342	0.00913	0.657
5	0.00292	0.292	0.00354	0.708
6	0.00141	0.254	0.00166	0.747
7	0.000765	0.225	0.000879	0.775
8	0.000451	0.202	0.000509	0.798
9	0.000283	0.183	0.000315	0.816
10	0.000187	0.168	0.000206	0.833
15	0.0000375	0.118	0.0000400	0.882

ing of the incompatibility system. Unless n is very small, the frequencies of individual singly and doubly heterozygous genotypes are very similar and are both close to the reciprocal of the total number of genotypes. This becomes apparent from equation (2) above, which, when $(n+1)^2 \gg 8$, gives $\alpha = 1$. The formulae for the relative frequencies of single and double heterozygotes has been used to deduce the expected probability of getting an incompatible reaction, when a single pollen type is used as a tester; this probability must depend on the number of loci and alleles (Lundqvist, 1962*b*), and could therefore be used as a method of estimating the number of alleles in a system with a known number of loci. The results presented here show that this purpose would be adequately served by simply assuming that all genotypes are present in equal frequency.

3. THREE INCOMPATIBILITY LOCI

A similar treatment is possible of three-locus systems. Writing p_1 , p_2 and p_3 for the frequencies of genotypes heterozygous at 1, 2 and 3 loci, respectively, the analogue to equations (1) is:

$$\begin{aligned} p_1 &= \frac{2}{W} \left(p_1 \frac{m_{11}}{2} + p_2 \frac{m_{21}}{4} + p_3 \frac{m_{31}}{8} \right) \\ p_2 &= \frac{2}{W} \left(p_1 \frac{m_{12}}{2} + p_2 \frac{m_{22}}{4} + p_3 \frac{m_{32}}{8} \right) \\ p_3 &= \frac{2}{W} \left(p_1 \frac{m_{13}}{2} + p_2 \frac{m_{23}}{4} + p_3 \frac{m_{33}}{8} \right) \end{aligned} \quad (3)$$

where $m_{i,j}$ is the number of genotypes that are heterozygous at i loci that would produce progeny heterozygous at j loci when pollinated by any particular gamete type, and which are compatible with pollen of that type. Arguments similar to those given above for two loci give the following m values:

$$\begin{aligned} m_{11} &= 3(n-1) - 1 \\ m_{12}, m_{13} &= 3(n-1) \\ m_{21} &= 6(n-1)^2 - 4(n-1) \\ m_{22} &= 6(n-1)^2 - 2 \\ m_{23} &= 6(n-1)^2 \\ m_{31} &= 4(n-1)^3 - 4(n-1)^2 \\ m_{32} &= 4(n-1)^3 - 4(n-1) \\ m_{33} &= 4(n-1)^3 - 4 \end{aligned}$$

No doubt larger numbers of loci could be dealt with, but this does not seem worthwhile in view of the restriction that equal numbers of alleles must be assumed for all loci.

Substitution of these expressions into equations (3) leads to a pair of simultaneous quadratic equations in the ratios p_1/p_2 and p_3/p_2 . Although

TABLE 2
Frequencies of the three classes of genotypes in the three-locus system, from equations (3)

Number of alleles at each locus (n)	Single heterozygotes		Double heterozygotes		Triple heterozygotes	
	Individual frequencies (p_1)	Aggregates	Individual frequencies (p_2)	Aggregates	Individual frequencies (p_3)	Aggregates
2	0.02406	0.289	0.04167	0.500	0.05283	0.211
3	0.00215	0.174	0.00295	0.477	0.00323	0.349
5	0.000107	0.080	0.000129	0.387	0.000133	0.533
10	0.00000181	0.024	0.00000199	0.242	0.00000201	0.733
15	0.000000164	0.012	0.000000175	0.174	0.000000176	0.815

apparently intractable analytically, these equations may be solved numerically for given values of n . From the form of the coefficients, it can be seen that, as n tends to infinity, equations (3) are dominated on the right-hand side by the contributions from p_3 ; furthermore, these contributions tend to equality with increasing n . It therefore follows that with increasing numbers of alleles, the frequencies of each class of heterozygote tend to equality, as in the two-locus case. Table 2 shows some numerical examples. These have been checked by computer calculations of the (very complex) incompatibility system with three loci.

4. DISCUSSION

The first conclusion that can be drawn from the results described above is that unless the number of alleles is very small, the frequencies of all individual genotypes are almost equal, so that the assumption that this is the case is, for most practical purposes, quite adequate. Tables 1 and 2 also illustrate the further conclusion that if there are many alleles, the aggregate frequency of double heterozygotes is larger than that of single heterozygotes in the two-locus system, and triple heterozygotes in the three-locus system are commoner than doubles. With a large number of alleles, the population consists overwhelmingly of the highest-order heterozygote. In studies of the genetics of incompatibility systems, there is always a possibility that a system that is in reality a two-locus system will be mistakenly thought to be a single-locus system (or, with more loci, that one or more will be missed), simply because no plant heterozygous at all the loci is included in the study (Lundqvist, 1975). The present results show that this possibility is often quite unlikely, and indeed multiple heterozygotes were found in very limited samples in several studies of multi-locus systems in grasses (*e.g.* Hayman, 1956; Murray, 1974), *Ranunculus acris* and *Beta vulgaris* (Lundqvist *et al.*, 1973; Østerbye, 1975, 1977; Larsen, 1977). The results on the frequencies of the single versus higher-order heterozygotes also suggest that, if it is acceptable to assume that there are equal numbers of alleles at all loci, a fairly good estimate of that number would be given by the frequency with which single heterozygotes are found (see tables 1 and 2). Obviously, this represents a great deal of work.

Another important result of the present approach to the population genetics of multi-locus incompatibility systems is that it shows clearly that

the basic equations for the genotype frequencies do not depend on the recombination fractions between the loci. This makes it seem unlikely that there could be any selection for tighter linkage between the loci. It has been suggested that the single-locus gametophytic systems may have evolved from multiple-locus systems by a process of eliminating recombination between the loci (*e.g.* Lundqvist, 1975). This does not seem likely, in view of the argument given above. This conclusion is supported by the results of computer calculations of a two-locus system with a modifier of the recombination fraction between them. These calculations yielded the same genotype frequencies as given in table 1 (thus incidentally checking equations (1)), and these frequencies were not changed when the recombination fraction was changed, nor did the recombination modifier change in frequency, but remained at whatever frequency it was initially assigned, except for transient changes due to the precise genotype frequencies in the initial population. Finally, this origin for the single-locus system is ruled out by the following considerations. Imagine a two-locus system in which a modifier reducing recombination has spread, such that there is now no recombination between the S and Z loci. Such a population will not behave in the same way as a single-locus system. The two-locus nature of the system will betray itself, for example, by the fact that double heterozygotes will show incompatibility with four types of pollen, not two. The only way that a single-locus system can be derived from one with multiple loci is by fixation of alleles at all the loci except one.

Acknowledgments.—I thank B. Charlesworth for helpful discussions, and Professor A. Lundqvist for reading and commenting on this manuscript.

5. REFERENCES

- HAYMAN, D. L. 1956. The genetical control of incompatibility in *Phalaris coerulea* Desf. *Aust. J. Biol. Sci.*, **9**, 321-331.
- LARSEN, K. 1977. Self-incompatibility in *Beta vulgaris* L. Four gametophytic, complementary S-loci in sugar-beet. *Hereditas*, **85**, 227-248.
- LUNDQVIST, A. 1954. Studies on self-sterility in rye, *Secale cereale* L. *Hereditas*, **40**, 278-294.
- LUNDQVIST, A. 1961. Self-incompatibility in *Festuca pratensis* Huds. *Hereditas*, **47**, 542-562.
- LUNDQVIST, A. 1962a. Self-incompatibility in diploid *Hordeum bulbosum* L. *Hereditas*, **48**, 138-152.
- LUNDQVIST, A. 1962b. The nature of the two-loci incompatibility system in grasses. II. Number of alleles at the incompatibility loci in *Festuca pratensis* Huds. *Hereditas*, **48**, 169-181.
- LUNDQVIST, A. 1965. Self-incompatibility in *Dactylis aschersoniana* Graebn. *Hereditas*, **54**, 70-87.
- LUNDQVIST, A. 1975. Complex self-incompatibility systems in angiosperms. *Proc. Roy. Soc. Lond. B.*, **188**, 235-245.
- LUNDQVIST, A., ØSTERBYE, U., LARSEN, K., AND LINDE-LAURSEN, I. 1973. Complex self-incompatibility systems in *Ranunculus acris* L. and *Beta vulgaris* L. *Hereditas*, **74**, 161-168.
- MURRAY, B. C. 1974. Breeding systems and floral biology in the genus *Briza*. *Hereditas*, **33**, 285-292.
- ØSTERBYE, U. 1975. Self-incompatibility in *Ranunculus acris* L. I. Genetic interpretation and evolutionary aspects. *Hereditas*, **80**, 91-112.
- ØSTERBYE, U. 1977. Self-incompatibility in *Ranunculus acris* L. II. Four S-loci in a German population. *Hereditas*, **87**, 173-178.